Tracking down social determinants of health in electronic health records

July 10 2023

Credit: Unsplash/CC0 Public Domain

Information on the nonmedical factors that influence health outcomes, known as social determinants of health, is often collected at medical appointments. But this information is frequently recorded as text within

the clinical notes written by physicians, nurses, social workers, and therapists.

Researchers from Regenstrief Institute and Indiana University Fairbanks School of Public Health recently published one of the first studies in which natural language processing was applied to social determinants of health. The researchers developed three new natural language processing algorithms to successfully extract information from text data related to housing challenges, financial stability and employment status from electronic health records.

"Health and well-being are not just about medical care. Mostly, they are about our behaviors, our environment, our social connections," said Regenstrief Institute Research Scientist and Fairbanks School of Public Health faculty member Joshua Vest, Ph.D., who led the study.

"More and more health care organizations are having to deal with social determinants because it is factors like financial resources, housing, and employment status that really drive costs that make people unhealthy. The challenge for health care organizations is effectively measuring and identifying patients with social risks so that they can intervene."

"Our work helps advance the field in both application and methodology. Natural language processing has been applied to numerous conditions in the past, but this is one of the first papers that applies it to social determinants of health."

"We demonstrated that a relatively simplistic natural language processing approach could effectively measure social determinants instead of using of more sophisticated deep learning and neural network models. These later models are powerful but complex, difficult to implement, and require a lot of expertise, which many health systems don't have."
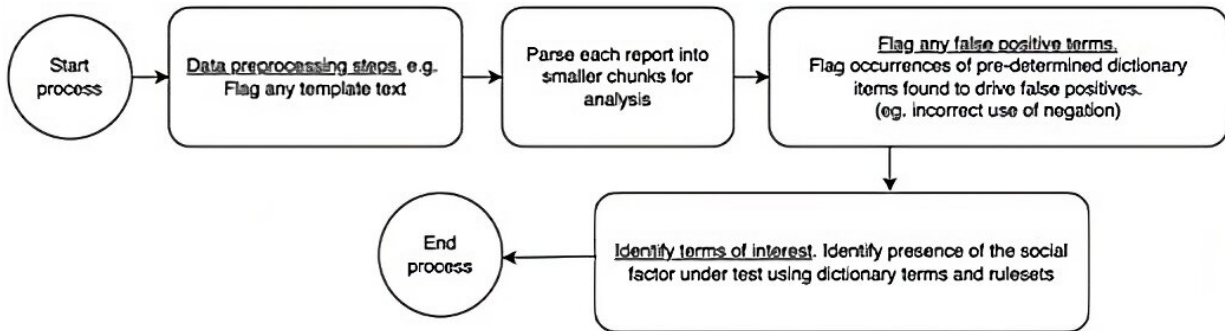
State machine process diagram. Credit: *JAMIA Open* (2023). DOI: 10.1093/jamiaopen/ooad024

Dr. Vest added, "We purposely designed a system that could run in the background, read all the notes and create tags or indicators that says this patient's record contains data suggesting possible concern about a social indicator related to health."

"Our overall goal is to measure social determinants well enough for researchers to develop risk models and for clinicians and health care systems to be able to use these factors—housing challenges, financial security and employment status—in routine practice to help individuals and to provide a better understanding of the overall characteristics and needs of their patient population."

Information indicating social needs can be extracted for many types of data in an electronic medical record, including information on patient occupation, health insurance coverage, marital status, size of household, address (low versus high crime area) and frequency of address changes.

Previously, Dr. Vest and colleagues, including Regenstrief Institute Vice President for Data and Analytics Shaun Grannis, M.D., created an app

they named [Uppstroms](#), Swedish for upstream, and successfully demonstrated that it could use structured data to predict patients in need of a referral to a social service such as a nutritionist.

The research is published in the journal *JAMIA Open*.

Provided by Regenstrief Institute